Name: Ali Bilal
Roll No: p20–0077
SECTION : 6C

In [78]:
```
#imports
from sklearn.datasets import load_iris
from sklearn import tree
import pandas as pd
```

In [79]:
```
data = load_iris()
```

In [80]:
```
type(data)
```

Out[80]:
```
sklearn.utils.Bunch
```

In [81]:
```
data = load_iris()
df = pd.DataFrame(data=data['data'], columns=data['feature_names'])
```

In [82]:
```
df.head(150)
```

Out[82]:

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
0	5.1	3.5	1.4	0.2
1	4.9	3.0	1.4	0.2
2	4.7	3.2	1.3	0.2
3	4.6	3.1	1.5	0.2
4	5.0	3.6	1.4	0.2
...	...	...	...	...
145	6.7	3.0	5.2	2.3
146	6.3	2.5	5.0	1.9

In [83]:
```
df.shape
```

Out[83]:
```
(150, 4)
```

In [84]:
```
df['target'] = data['target']
```

In [85]:
```
print(df.head())
```

```
   sepal length (cm)  sepal width (cm)  petal length (cm)  petal width (cm)  \
0                5.1               3.5                1.4               0.2
1                4.9               3.0                1.4               0.2
2                4.7               3.2                1.3               0.2
3                4.6               3.1                1.5               0.2
4                5.0               3.6                1.4               0.2

   target
0       0
1       0
2       0
3       0
4       0
```

In [96]:
```
X = df[['sepal length (cm)', 'sepal width (cm)', 'petal length (cm)', 'petal width (cm)']]
y = df['target']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=0)
```

In [97]:
```
print(X_train)
print(X_test)

print(y_train)
print(y_test)
```

```
73     1
54     1
63     1
```

In [98]:
```
data.target_names
```

Out[98]:
```
array(['setosa', 'versicolor', 'virginica'], dtype='<U10')
```

In [89]:
```
DTC_Model =DecisionTreeClassifier(criterion= 'entropy', max_depth=4, random_state=10)

DTC_Model.fit(X_train, y_train)
```

Out[89]:
```
DecisionTreeClassifier(criterion='entropy', max_depth=4, random_state=10)
```

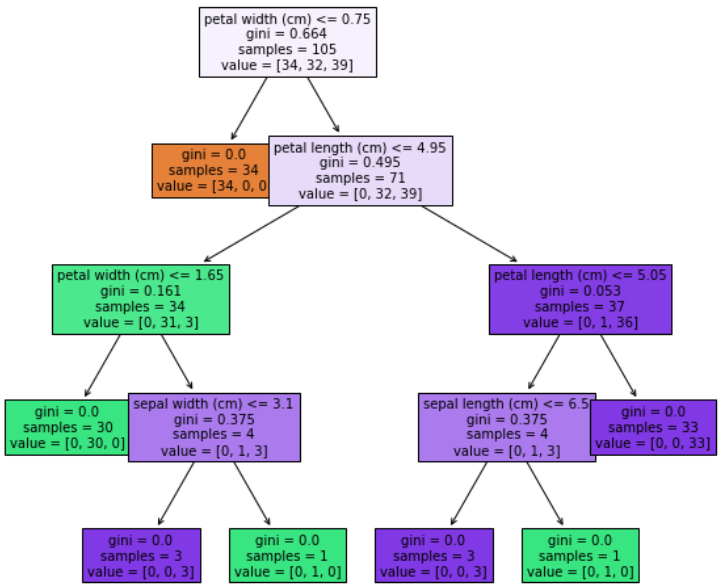
In [90]:
```
DTC_Model.score(X_train , y_train)
```

Out[90]:
```
1.0
```

In [91]:
```
y_pred = DTC_Model.predict(X_test)
```

In [102]:
```
# DTC_tree = tree.plot_tree(DTC_Model)
plt.figure(figsize=(10, 9))

DTC_tree = tree.plot_tree(DTC_Model, filled=True,
                          feature_names=['sepal length (cm)', 'sepal width (cm)', 'petal length (cm)', 'petal width
                          fontsize=10,
                          )
```



In [103]:
```
#imports
from sklearn.datasets import load_iris
from sklearn import tree
import pandas as pd
```

In [104]:
```
data = load_iris()
```

In [105]:
```
data = load_iris()
df = pd.DataFrame(data=data['data'], columns=data['feature_names'])
```

In [106]:
```
df['target'] = data['target']
```

In [107]:
```
X = df[['sepal length (cm)', 'sepal width (cm)', 'petal length (cm)', 'petal width (cm)']]
y = df['target']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=0)
```

In [108]:
```
DTC_Model =DecisionTreeClassifier(criterion= 'gini', max_depth=4, random_state=10)

DTC_Model.fit(X_train, y_train)
```

Out[108]:
```
DecisionTreeClassifier(max_depth=4, random_state=10)
```

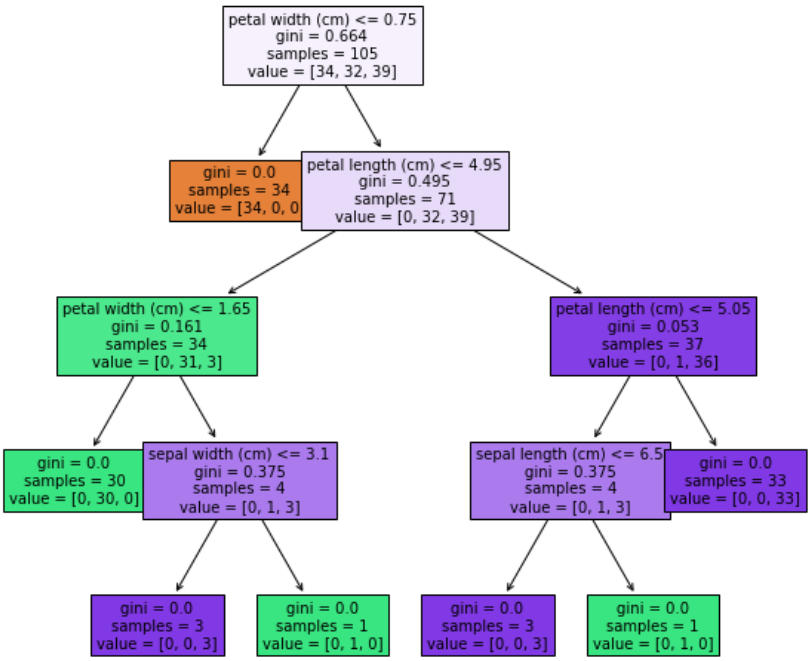
In [109]:
```
DTC_Model.score(X_train , y_train)
```

Out[109]:
```
1.0
```

In [110]:
```
y_pred = DTC_Model.predict(X_test)
```

In [113]:
```
# DTC_tree = tree.plot_tree(DTC_Model)
plt.figure(figsize=(10, 9))

DTC_tree = tree.plot_tree(DTC_Model, filled=True,
                          feature_names=['sepal length (cm)', 'sepal width (cm)', 'petal length (cm)', 'petal width
                          fontsize=10,
                          )
```


petal width (cm) <= 0.75
gini = 0.664
samples = 105
value = [34, 32, 39]
gini = 0.0
samples = 34
value = [34, 0, 0
petal length (cm) <= 4.95
gini = 0.495
samples = 71
value = [0, 32, 39]
petal width (cm) <= 1.65
gini = 0.161
samples = 34
value = [0, 31, 3]
petal length (cm) <= 5.05
gini = 0.053
samples = 37
value = [0, 1, 36]
gini = 0.0
samples = 30
value = [0, 30, 0]
sepal width (cm) <= 3.1
gini = 0.375
samples = 4
value = [0, 1, 3]
sepal length (cm) <= 6.5
gini = 0.375
samples = 4
value = [0, 1, 3]
gini = 0.0
samples = 33
value = [0, 0, 33]
gini = 0.0
samples = 3
value = [0, 0, 3]
gini = 0.0
samples = 1
value = [0, 1, 0]
gini = 0.0
samples = 3
value = [0, 0, 3]
gini = 0.0
samples = 1
value = [0, 1, 0]